Recall Fluency: a Cognitive Source of Belief and Choice Heterogeneity

Pedro Bordalo, Nicola Gennaioli, Matthew Obrien, Andrei Shleifer

March 2026. Preliminary, do not circulate

Abstract

A growing literature shows that beliefs reflect cued recall from a database, subject to interference across experiences. We test this model in the HRS dataset which is unique in offering a respondent-level proxy for interference based on recall fluency in the word-list task, a standard paradigm in memory research. Examining beliefs across three domains – stock returns, home price growth, and life expectancy – we show expectations satisfy three key predictions. People with low interference (i.e. high recall fluency) exhibit: 1) more sensitivity to past experiences, both domain-specific and non domain-specific, 2) more sensitivity to current cues, and 3) less idiosyncratic heterogeneity. This implies that people with low interference are on average more optimistic about stocks, houses, and life expectancy, which correlates with higher stock and housing investment and a lower likelihood of retirement.

1. Introduction

A growing literature shows that regularities in human memory unify well documented features of measured beliefs, in particular distortions in the way beliefs react to news or current events (which reflects the role of cues), and the influence of remote DS and NDS personal experiences (which are due to the memory database and interference). Experimental and field evidence shows that exogenous changes in cues or the database of experiences affects recall, reasoning about forecasts, and the forecasts themselves, with qualitatively similar effects across abstract probability judgments, familiar economic domains, or assessment of novel risks.

So far, this work has explored variation in the cues that people are exposed to or in their past experiences, but not in their recall faculties. It is of course a fact of life that some people exhibit better memory than others. This variation is also robustly documented in abstract tasks in which people must recall familiar words from a list. Does variation in such recall fluency influence the way in which people form beliefs? Does a general-purpose feature of a person's mnemonic apparatus affect beliefs through the well-established mechanisms of cues and database across different domains? Addressing this question is important because variation in recall fluency could be an important source of belief and then choice heterogeneity across people, affecting economic outcomes. Just as people who are less patient save less and hence become less wealthy, so the beliefs of people with lower recall fluency may exhibit a

systematically different reaction to news or experiences, shaping their choices and market outcomes.

In this paper we address this issue by suing the HRS database a bi-annual survey of a representative panel of US residents who are 50 years or older, with about 20 thousand people per wave. The HRS has measured beliefs in different domains, including future stock returns, home price growth and life expectancy for many years, which allows to study both cross sectional heterogeneity and reaction to events over time, including major shocks such as the 2008 financial crisis. It also allows to compute, for each respondent, a rich database of experiences, which may be domain-specific (DS) for some beliefs but non domain-specific (NDS) for others. Crucially, and uniquely among such surveys, it measures respondents' recall fluency in abstract word tasks, a standard paradigm of measurement and theory in the literature on memory, starting from Ebbinghaus (1885). We show that recall fluency declines somewhat with age but is a surprisingly stable trait that varies across people and exhibits large orthogonal variation with respect to other cognitive traits such as IQ.

Consistent with the memory literature, we take recall fluency to capture, at least in part, an inverse proxy for memory interference: people may be less able to recall familiar words from a recently studied list because many other words they know appear superficially similar and hence block retrieval of the target words in the list (just like confusion about a specific event often reflects the inability to separate it from many other related events). In this view, recall fluency captures a key force of mnemonic beliefs models, which allows us to use these models to make predictions on how individual level heterogeneity in recall fluency should shape beliefs. Specifically, for people with high recall fluency, beliefs should: 1) be more sensitivity to past

experiences, both DS and NDS, 2) be more sensitivity to current events or cues, and 3) display less idiosyncratic heterogeneity.

Crucially, since recall fluency is a general-purpose technology, these predictions hold for a range of domains – stock returns, housing and life expectancy – that are unrelated to the word task where fluency is measured. Predictability from recall fluency emphasizes the role of memory in a robust way (compared to a link between beliefs and recollections, which may be due to reverse causality).

We find strong empirical support for these predictions. Beliefs reflect past experiences to an extent that increases in RF: experience effects are not mechanical, rather recalling them is key. This complements earlier results of similarity- or cue-dependent experiences effects, but identifies a persistent component. We also show that reaction to news stronger if RF higher and the news "matches" the personal experience database, that is, if they are particularly similar to a past experience. Thus, beliefs reflect not mechanical distortions in the use of statistics or experiences, but rather the ability to remember, and imagine, events the news are similar to. Variation in recall fluency generates strong heterogeneity in beliefs and instability in the reaction to news. Critically, variation in RF and in experiences, through their effects on beliefs, predict economically important choices such as investment in stocks, home purchases, and retirement decisions.

We break new ground by connecting beliefs and experiences to a measurement of general-purpose memory technology: recall fluency in standard laboratory memory tasks.

Memory capacity emerges as a trait – like risk aversion, patience, or optimism – that can shape beliefs and choice.

2. Motivating Facts

2.1 The HRS survey

The Sample. The University of Michigan Health and Retirement Study (HRS) is a nationally representative biennial panel of U.S. adults over age 50. It introduces a new cohort roughly every six years and follows respondents (and their partners) until death. We consider ten waves, starting in 2002 because our main belief variables are not collected before. We have 34,677 unique respondents and 189,097 person-wave observations. Not all respondents appear in every wave due to staggered entry, mortality, attrition. Appendix Table X describes the survey and our sample. Our respondents have on average 67.7 years for age (with 11.4 st dev), 12.5 years of education (3,2 st dev), income of $57,574 ($105,371 st dev),[1] 53% of people are retired, 63% are married, 38% female.

Outcomes: Beliefs and Choices. The HRS elicits beliefs over various social and personal events. We study two macroeconomic domains, the stock and housing markets, and a personal domain, life expectancy. Beliefs about these domains play an important role in economic models, they are available in many waves, and are easily traceable to personal experiences.[2]

[1] In this variable we combine all income streams: salaried income, self-employed income, bonuses, social security benefits, workers compensation, and pension. The combination of different income sources creates a high variance. For instance, just the 'salaried income' has a standard deviation of only 60k.

[2] For instance, the HRS elicits also expectations about inflation and US economic performance, but only in a few waves (1996, 1998, 2000). It also elicits a range of personal outcomes which cannot however easily trace to experienced measured within the HRS survey (e.g. beliefs about inheritance or financial help from family members).

The stock market question is “What is the percent chance that the stock market will be worth more in one year than today?”. The housing question is “What is the percent chance that houses will be worth more in one year than today?”. The life expectancy question is: “What is the percent chance that you will live until the age of 80/85/90/95/100?”, where the threshold varies with respondents’ age: it is 80 for people under 70, 90 for people between 75 and 80, 95 for people between 80 and 85, and 100 for people between 85 and 90.

The HRS also measures a broad range of choices. To see if beliefs matter, we consider 6 choices tied to our domains of interest. For the stock market, we consider whether or not the respondent has made a new stock market investment in the past 2 years (regardless of amount), and the percentage of the respondent’s IRA that is invested in stocks or mutual funds, which is restricted only to respondents who have an IRA. For the housing market, we consider whether the respondent has bought a new home in the past 2 years. For health choices, we consider whether the respondents expects to work past the age of 70 (restricted to respondents who are currently working), whether the respondent is currently retired or not, and the percentage of a respondents income invested in their pension (restricted to respondents on a defined contribution plan).

Demographics. The HRS report a vast array of indicators measuring respondents’ cognitive faculties, health, financial and personal experiences, as well as standard demographics. We use age, gender, marital status, retired/not, income, and education as baseline controls in our regressions.

Experiences. Measured experiences comprise macroeconomic outcomes lived by respondents as well as personal economic and health outcomes. For macro experiences, we focus on experiences directly tied to beliefs.

Stock market and Housing experiences. We generate a respondent's database of annual real stock market returns over their lifetime (birth to 1-year ago), constructed from the Shiller S&P500 dataset (https://shillerdata.com/). In the housing market, we likewise measure housing returns from the Federal Housing and Finance Agency's (FHFA) home price index (https://www.fhfa.gov/data/hpi/datasets?tab=quarterly-data), which is estimated using sale and appraisal data. We match the respondents census region of birth – our most granular measure of location in the survey data – to the respective regional FHFA housing index. Unlike in the stock market, housing price data is only available from 1975 onwards. Hence, our measure of a respondent's housing market database consist of annual real housing market returns in the respondent's region of birth, spanning 1975 to present. These experiences are Domain Specific (DS) for stock and housing market beliefs, they are Non-Domain Specific (NDS) for life expectancy.

Health Experiences. To generate analogous health Domain Specific experiences, we rely on the RHS dataset directly. In particular, we start by constructing measures of the following: i) childhood health and ii) current health. Childhood health is the first principal component of the following variables: missed time in school for health reasons (1,0), disabled as a child (1,0), number of childhood health conditions, number of severe childhood health conditions, and number of psychological childhood health conditions. Current health is the first principle components of the following variables: days spent in bed last month, number of health conditions, self-rated health (1-5), depressed (1,0), recent heart attack (1,0), new/worsening

severe health condition (1,0), sleep PCA, number of daily tasks the respondents has difficulty doing on their own, whether the respondent has a functional helper (1,0), and number of tasks a functional helper helps with on a daily basis). We then generate a respondent's database of health experiences as the difference between health levels across surveys. In the respondent's first survey, we use the difference between health in the current survey, and childhood health.

Personal experiences. We use i) childhood health and ii) childhood socioeconomic status. Childhood health is as described above, and childhood socioeconomic status is the first principal component of the following variables: father's years of education, mother's years of education, father was unemployed as child (1,0), moved due to financial difficulty as a child (1,0), family received financial aid as a child (1,0), lived rural (1,0), spoke English at home (1,0), whether the respondent's mother worked as (1,0). These experiences, which are measured in the HRS, are NDS events for the stock and housing markets, but they can be DS for life expectancy (i.e. past personal health). We later discuss DS, NDS, and our model's predictions about them.

2.2 Recall Fluency and IQ

The HRS also measures key cognitive faculties, specifically memory and mathematical ability. Memory is our central focus, and in Section 2.3 we link both measures to beliefs.

Memory is measured using three tasks. In two tasks respondents first study a list of 10 words. Then, they are asked to list as many of these words as possible: i) right after the study period (Immediate Memory task), and ii) after completing a five minutes questionnaire on mental health (Delayed Memory task). Memory performance is then measured as successful recall of words in the list. In the third task, which is available in waves 2010-2020, respondents must

name as many animals as possible in 60 seconds. Here memory performance is the number of animals recalled.

Starting from Ebbinghaus (1883), free recall of words from a list is a canonical tool for studying human memory in cognitive science. The advantage of this method is that it allows to study general purpose regularities in human recall in an almost idealized context. It abstracts from differences in education and world knowledge (the words are common and hence familiar to subjects), and does not entail symbol manipulation or logical rules, so it does not require reasoning. Based on these tasks, a vast body of work has unveiled four robust recall patterns.

The first is primacy: words studied earlier in the list are more likely to be recalled, due to repeated rehearsal. This phenomenon reflects a general force: recall of an event increases in its experienced frequency. The other patterns are: i) recency (words studied later in the list are more likely to be recalled compared to words in the middle), ii) contiguity (after recalling a word, a person is more likely to recall a contiguous word in the list), and iii) semantic proximity (after recalling a word, a person is more likely to recall a related word, e.g. car -> brake).

The last three patterns point to the importance of similarity along different features. Recency captures the similarity of mental context between the time in which the word was studied and the subsequent recall phase. Contiguity captures the similarity of mental context when the recalled and the contiguous words were studied. Semantic proximity captures similarity or relatedness of the words' meanings. Besides frequency and similarity, a key additional force is interference. It refers to the fact that cues prompting recall of specific word also inhibit, due to memory limits, the recall of words on the list that are less similar to the cue. Interference also

arises when spontaneous thoughts cue words outside of the list, which may be erroneously recalled (the so called extra-list intrusions) prompting forgetting of words in the list.

Existing work shows that frequency similarity and interference do not operate uniformly. Performance in word recall tasks is shown to vary within a person across different tests and across people on average, with some subjects exhibiting higher performance than others. A variance decomposition exercise shows that cross-sectional variance in memory accounts for about two-thirds of the variation, with within person variation over time accounting for the rest. Inter-personal variation of memory performance is our key focus here.

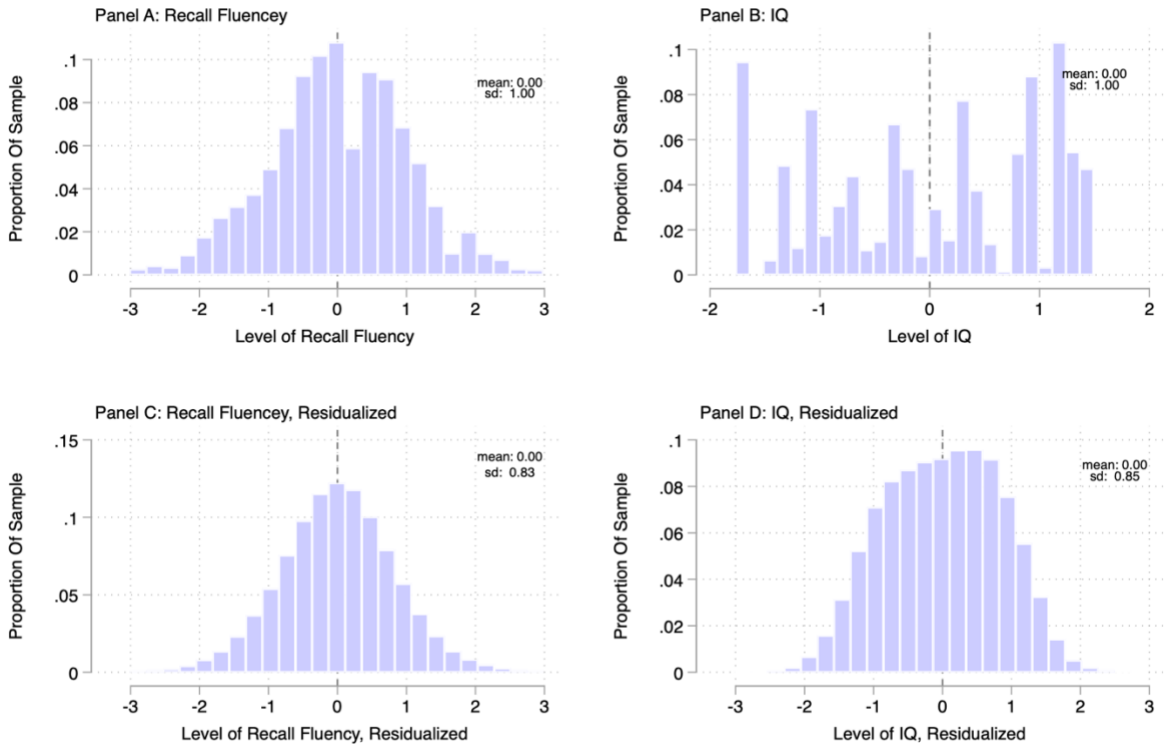
Mathematical ability is also measured in the HRS using three tasks. The first task, counting backwards, requires respondents to perform a sequence of subtractions starting from a given number. In the second task, number series, respondents must guess a rule by filling the last item of a series. In the third task, numeracy, respondents must convert percentages into absolute frequencies. Higher performance in each task indicates higher math ability. Counting backward is available in all years, number series is elicited in the 2010, 2012, 2016 and 2020 waves, and numeracy is only asked to first-time respondents. Mathematical ability affects the use of statistical information, a key driver of beliefs in standard models, also if someone recall every bit of information from memory. Controlling for mathematical ability therefore allows us to separate memory – our main object of interest – from this confound.[3]

[3] Additional measures include: 1) General Knowledge, which are a series of questions like "Who is the president?", "what is today's date", etc. 2) Word Meanings, which asks what the meanings are of words like "repair", "fabric", "plagiarize", 3) Verbal analogies. For example, "Mother is to Daughter as Father is to [blank]". These measures are available for much smaller samples, e.g. general knowledge is asked to half the respondents, word meanings are asked only to first time respondents, etc.

To capture memory and mathematical ability at the respondent level, we construct two indices. The first one increases in memory performance and we call it called *Recall Fluency* (RF). The second index increases in mathematical performance. We use the shorthand "IQ" given the strong positive association between mathematical ability and this measure of intelligence. Each index is constructed by standardizing individual memory or math components across waves and by averaging, in each wave, a respondent's available standardized components.

Figure 1 reports the raw distribution of RF in Panel A, that of IQ in Panel B. It then reports in Panel C the residualized distribution of RF after regressing it on IQ and a set of demographics (age, income, education, marital status, retired/not, gender), reports in Panel D the that of IQ after regressing it on RF and the same demographics.

Figure 1. Distribution of Recall Fluency and IQ

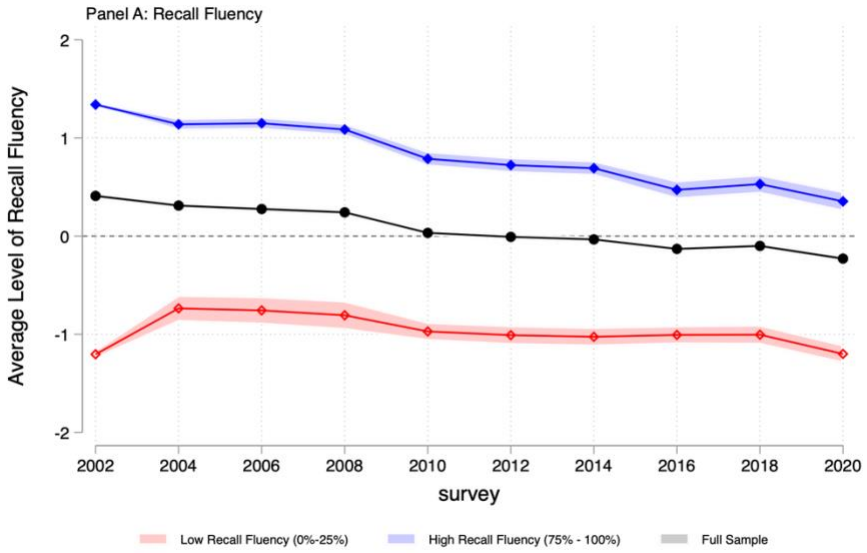


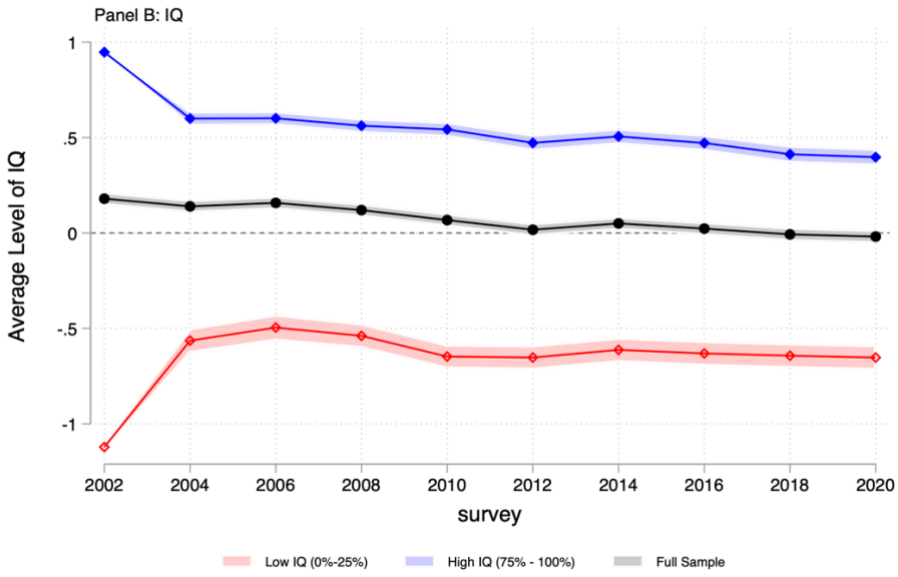
There is significant variation in RF and IQ across people and survey waves. The two indices are positively associated, their correlation is 0.4. Memory tasks do not require any mathematical ability, but some of the math tasks (e.g. counting backwards) benefit from strong memory. More broadly, there is a positive association between different dimensions of cognitive prowess, see the Appendix. Critically, though, Panels C and D show that even after controlling for demographics and for the association between RF and IQ, orthogonal variation in the proxies is large: the standard deviation of RF shrinks only by 0.17, that in IQ by 0.15. In terms of explanatory power, the R2 from a regression of RF on demographics is 0.24, upt to 0.3 is we include IQ; the R2 from a regression of IQ on demographics is 0.21, up to 0.28 if we include RF. The large independent variation allows us to study the role of RF when controlling for many confounds,

including IQ. We also run non-parametric regressions, across RF and IQ quintiles, to allow for general dependencies.

Figure 1 reflects both stable interpersonal differences and within person variability. Is inter-personal variation significant? To what extent does it reflect age-dependent decay in cognitive faculties? Figure 2 show cross sectional and over time differences in RF and IQ. It focuses on the subset of 3,643 respondents present in the totality of our waves and reports the time evolution the index's average among all respondents (black line), among respondents who were in the top 25% of the index in 2002 (blue line) and among those who were in the bottom 25% in 2002 (red line). Panel A reports the results for RF, Panel B for IQ.

Figure 2. Cross Sectional vs Over Time Variation in RF and IQ





There is over time memory decay: average RF declines from 0.38 in 2002 to -0.20 in 2020, amounting to 0.58 standard deviations. Critically, the blue and red lines reveal large and stable interpersonal differenced in RF. In 2004 the top quartile has RF of 1.03 and the bottom quartile of -0.65, a difference equal to 1.68 standard deviations.[4] In 2020 the difference between the two groups amounts to the comparable 1.4 standard deviation. Similar patterns appear, less pronounced, for IQ: the is a slight average decay from 0.18 in 2002 to -0.02 in 2020, amounting to 0.2 standard deviation, but stable cross sectional-gaps.

In sum, there is significant variation in how HRS respondents perform in standard word recall tasks, which have been historically used to unveil key regularities in human memory. RF is

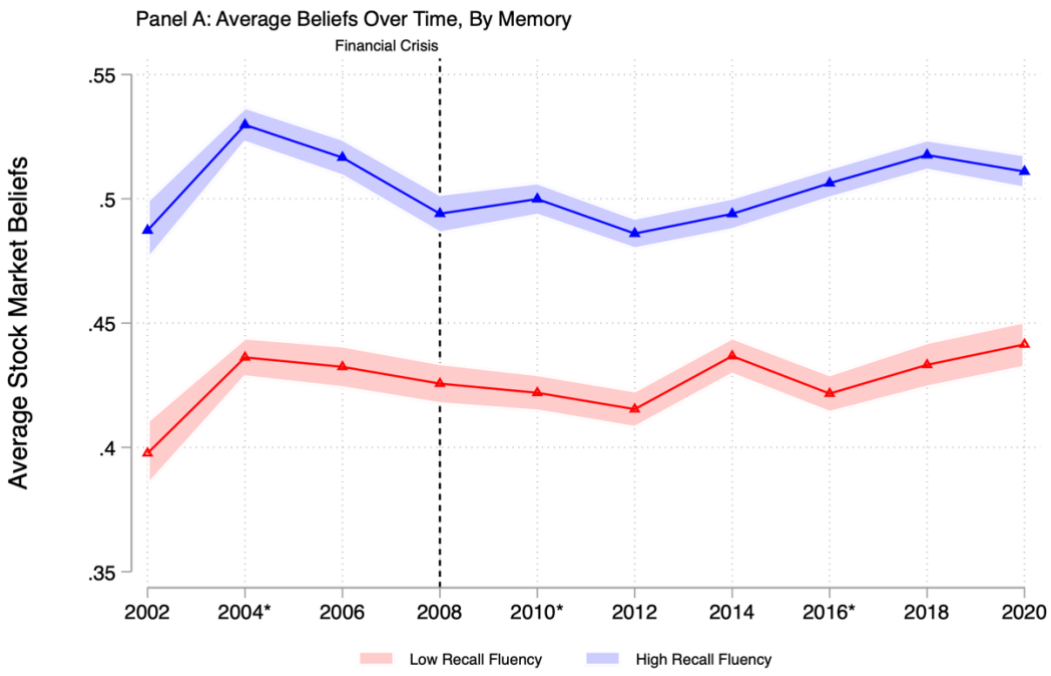
[4] We cannot compare quartiles in 2002 because here variation in RF is also attributable to random shocks, which revert in 2004. A different way to quantify differences is to note that the "within" standard deviation due to age is 0.59 while the "between" standard deviation due to average differences is 0.68, implying that the latter accounts for 54% of the total variance in RF. If we residualize respondents' average RF by average IQ we see that the standard deviation drops by 0.13 standard deviations, so there is significant orthogonal variation in the two measures.

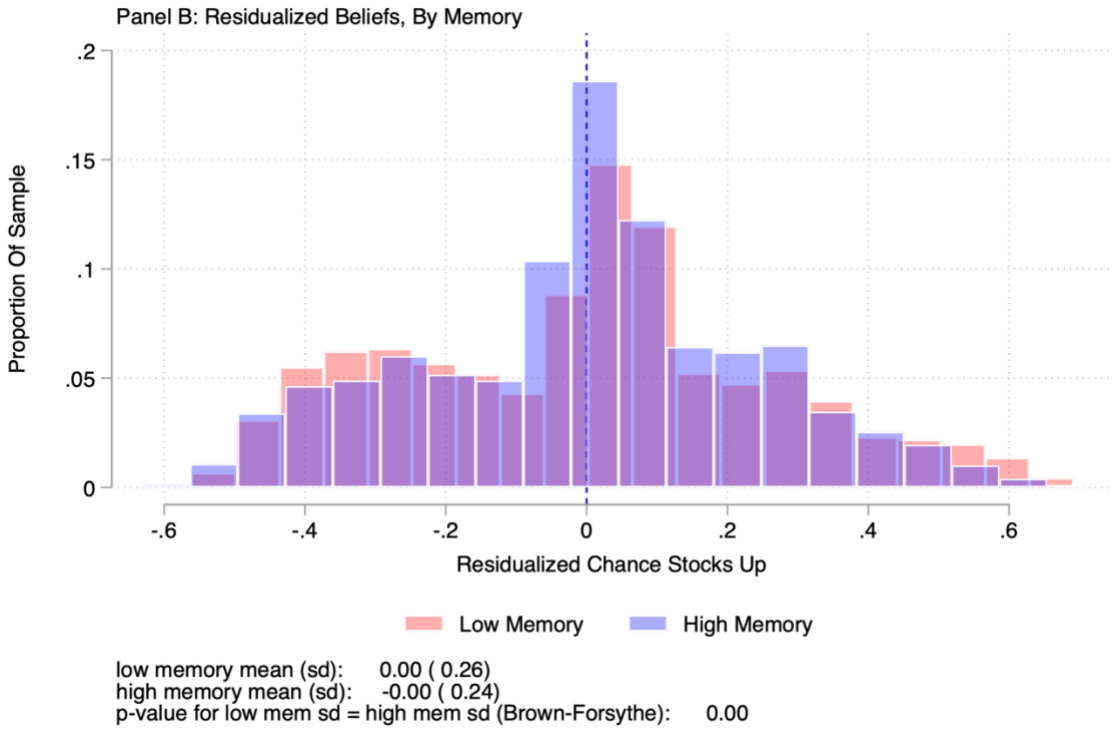
therefore a stable personal trait. We now show that such trait plays an important role in belief formation about aggregate and personal events, and – as we will see – in economic behavior.

2.3 Recall Fluency and Beliefs

We report descriptive facts about the link between beliefs and the RF index. Figure 3 reports in Panel A the average estimate for the "percent chance stocks are worth more in 1year" across survey waves for the top (blue) and bottom (red) RF quartile. In Panel B it reports the overall distribution of estimates for the high and low memory groups, residualized by IQ, age, income, education, gender, retired and marital status.

Figure 3. Average Stock Market Beliefs Over Time, by Memory





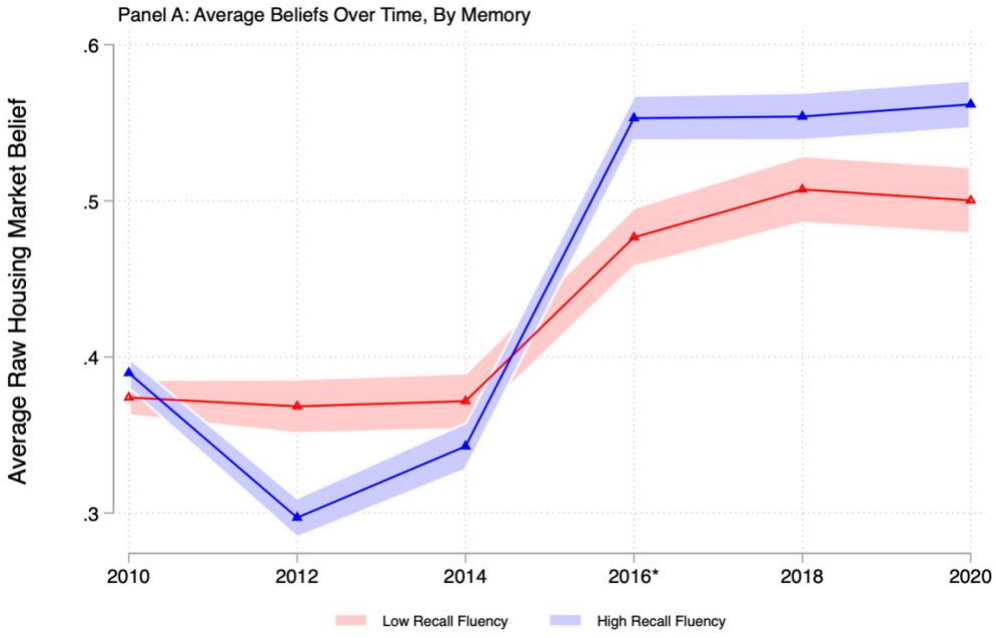
There is a sizable difference between high and low RF people: the former estimate an 8% higher probability of a stock market rise compared to the latter. We also see a difference in how these groups react to the 2008 crisis (dashed vertical line): the high memory group feature sharper boom-bust beliefs: they become somewhat more optimistic from 2002 to 2004 compared to low memory (probability increases by 4.2pp and 3.8pp, respectively), they exhibit a sharper drop during the crisis (probability decreases from peak by 2.3pp and 0.6pp, respectively), and exhibit a larger recovery from through to 2018 (probability increases by 2.3pp and 0.7pp, respectively).

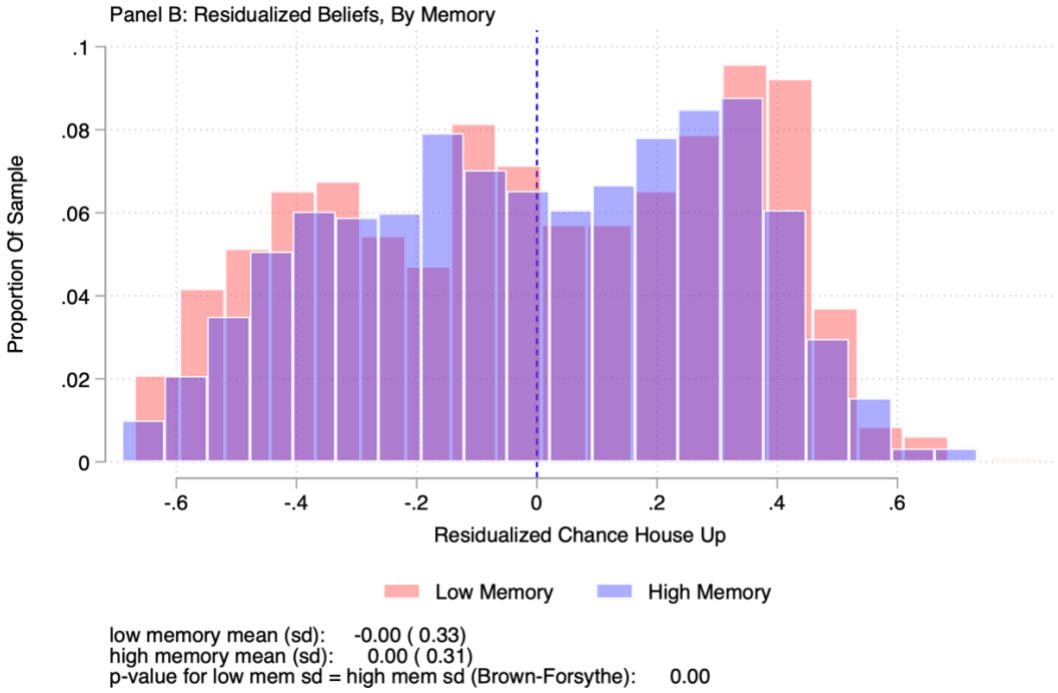
Both groups are on average too pessimistic about stock return: the average estimate of high memory people falls short of the actual frequency of stock market increase in the sample by

31 percentage points and the low memory group by 38 percentage points. The two groups exhibit similar distributions but high memory has lower belief dispersion after controlling for covariates.

Figure 4 reports in Panel A the average estimated "percent chance of home worth more in 1year compared to today" for the top (blue) and bottom (red) RF quartiles, and in Panel B the distribution of individual estimates for the two groups after residualizing based on the same variables as in Figure 3. This question is asked to people who a) fully own, or b) are paying off the mortgage on their home, corresponding to 44% of the sample. The question is only available from 2010 onwards (so, after the 2008 crisis).

Figure 4. Average Housing Market Beliefs Over Time, by Memory





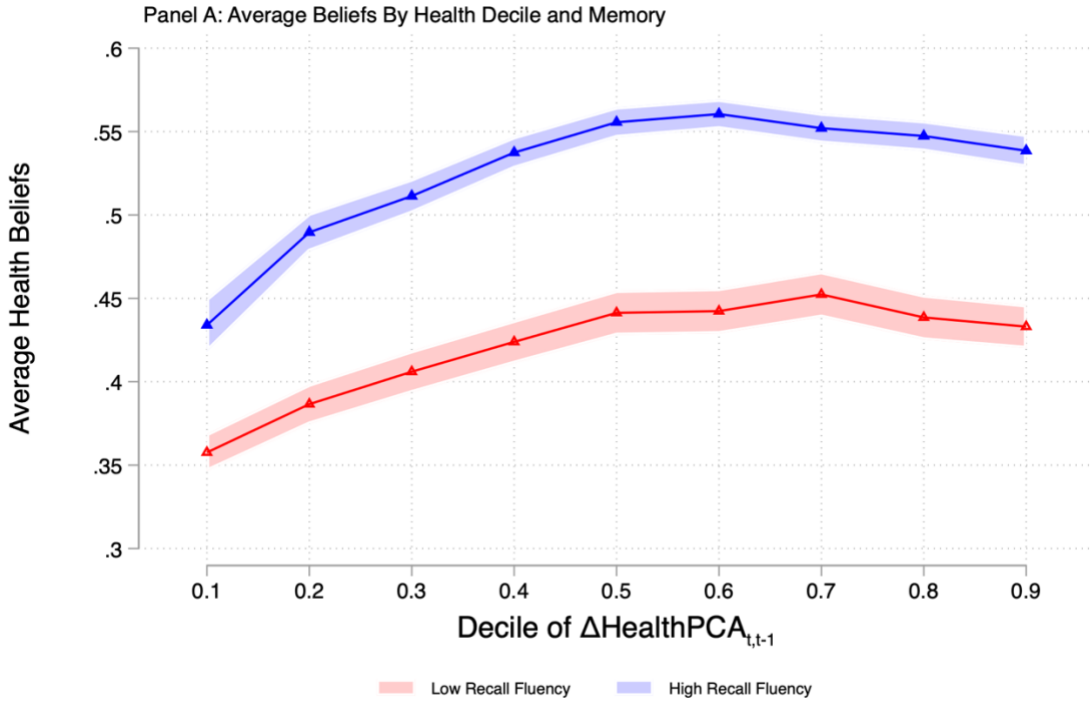
Again, there are significant differences tied to memory. Here the "level" of beliefs is comparable but again the high RF people have more volatile housing beliefs than low RF ones: they drop more after the crisis (drop of 1.5pp vs 0.5pp), when the housing market busts, but recover faster afterwards (recovery of 26.5pp vs 13.2pp).

Both groups on average underestimate the probability of house price increases compared to their actual frequency (by 28.2pp for high memory and by 29.7pp for low memory). The cross-sectional distribution of beliefs in each group is again similar and features fat tails, but the belief dispersion for low memory (33pp) is greater than that of high memory (31pp), after controlling for covariates.

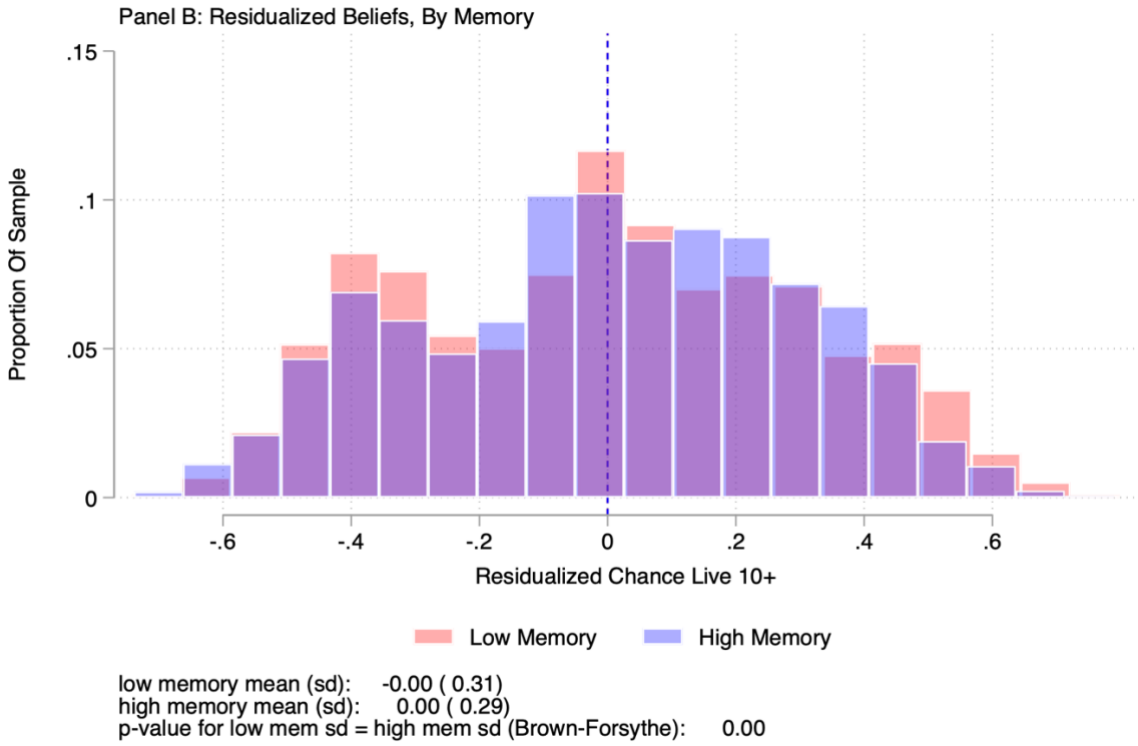
Finally, consider life expectancy, a personal event. Figure 5 reports in Panel A the average estimated probability of living 10 more years for the top (blue) and bottom (red) RF quartile as a

function of a respondent's current health condition, proxied by the change in the first principal component of health status proxies, as described above.[5] In Panel B we report the cross sectional distribution of beliefs computed using the previous proxy but now also controlling for the current health PCA measure.

Figure 5. Average Health Beliefs By Decile of Health Changes



[5] First principal component of the following variables: days spent in bed last month, # total health conditions, self-rated health, depressed, recent heart attack, New/Worsening severe health condition, sleep PCA, total number of task difficulties, has a functional helper, tasks helped with on a daily basis.



High RF people are significantly more optimistic about their life expectancy than low RF people at any current health status, and their optimism also grows faster as health status improves. The cumulative change in the blue curve (high RF) is 10.5pp, larger than the cumulative change in the red curve (low RF) of 7.5pp. Similarly to stocks and housing, low memory people exhibit also greater belief dispersion than high memory ones, 29pp for high RF compared to 31pp for low RF, after controlling for demographics.

These findings are surprising: individual performance in abstract laboratory word recall tasks is associated with sharp differences in beliefs about macroeconomic and personal events. RF can account for sizable baseline belief heterogeneity across people in terms of mean and standard deviation of opinions, but it also shapes their differential reaction to news. Why is this the case? A standard answer is omitted variables. High RF people have different information or

preferences than low RF people. This possibility is not entirely plausible because, as we said, word recall tasks do not rely on knowledge or preferences and it is largely independent of demographics proxying for standard drivers of beliefs. Another possibility is that variation in RF captures variation in memory, which in turn causes differences in beliefs. But this possibility raises many difficult questions: why does RF affect beliefs in domains that are so different? Does it capture a parameter of a general purpose belief technology? And why does it affect beliefs and reaction to news so differently across domains? To understand the role of RF and to try to address these questions, we now introduce a model of belief formation under selective memory.

In this model, memory is a general technology affecting beliefs in real world domains, and – in an analogy with production functions in economics – the parameters of this technology shape both performance in laboratory recall tasks, and hence RF, and ability to recall data from memory in the field, and hence beliefs. The model yields predictions that allow to: i) test memory against competing explanations, and ii) shed light on the differential role of RF in different domains.

3. The Model

We formalize recall based on memory research (Baddeley 199?, Kahana 2012), and link between memory to beliefs following the recall and simulation model of Bordalo et al (2023).

3.1 The Recall Technology

As in standard memory models, we rely on: i) a database of experiences, ii) the memory cue, and iii) a retrieval function mapping the cue to recollections from the database.

The database is a frequency distribution over categories of experiences. There is a mnemonic space E. A category is a trace/vector $e \in E$. Let f_e be the number of lived instances of $e \in E$; The database is $\mathbf{f} = \{f_e\}_{e \in E}$. Vector e could report the objective description/signal $s \subset \mathbb{R}$ of the experience, its subjective valence $v \subset \mathbb{R}$, and its semantic context $c \subset C$, where C is a context space, so that $e = (s, v, c)$. A stock return experience would then report a numerical return range, say $s = (2\%, 3\%)$, a marker for positive emotional valence, say $v = \mathbb{R}_{++}$, and context features such as the semantic "stock market" domain, the year, the country, etc. The database specifies the frequency of lived episodes falling in this and in all other categories.

Similarity and the Cue. The mnemonic space is endowed with a function $d: E \times E \to \mathbb{R}$ measuring the discrepancy between any two experiences $e, e' \in E$, where $d(e, e')$ increases in the difference among their features. If e' is a personal illness, it differs from a stock market experience along the objective, subjective, and context aspects, entailing a large $d(e, e')$. We shall not explicitly carry the features of experiences around, but it is useful to keep them in mind to follow the model's intuition for recall and belief formation.[6]

A cue is a subset of the mnemonic space $q \subseteq E$ that ignites similarity-based retrieval from it. The similarity $S(e, q)$ between the cue q and experience e is a bounded function that decays in their discrepancy $d(e, q)$. The probability of recalling e based on q from database $\mathbf{f}$ obeys:

$$r_e(q, \mathbf{f}) = \frac{S(e, q) \cdot f_e}{\sum_{e' \in E} S(e', q) \cdot f_{e'}}. \quad (1)$$

[6] Rather than as a distribution over a metric space E (Kahana 2012), distributed memory models formalize the database it as a sum of experience vectors (Murdock). These approaches have similar implications for our analysis.

The numerator captures similarity and frequency: experiences more similar to the cue and more frequently lived are more likely to be recalled. The denominator captures interference: even if e is similar to q and frequent, its recall can be blocked by an even more similar to q or frequent experience $e' \neq e$.

3.2 The Recall Technology and RF

We now link Equation (1) to the RF index, which helps us identify sources of variation in RF across people. This in turn, will give us new predictions about the memory mechanism that that we test using the HRS expectations data. Our analysis here is heuristic but it could be precisely formalized using contextual drift models.

Study phase. Study of a word list shapes the database. A word in serial position $j = 1, \dots, n$ entails an objective signal/phonetic w_j and a context tag c_j capturing the word's semantics and temporal features (we abstract from valence here but introduce when we study beliefs). Let $L \subset E$ be the set of traces (w_j, c_j) created in a word list task. Their encoding and rehearsal during study changes the database, which now attaches positive frequency f_j to each e_j.

Recall phase. The initial cue q_0 is the word list $(w_j)_j$ itself and the initial context c_0. As words are recalled, they shape the cue and subsequent retrieval. Primacy effects arise because the first word studied w_1 tends to be rehearsed more often than other words, increasing its encoded frequency f_1 and thus its recall in Equation (1). Recency effects occur because the last word studied w_n has a temporal context c_n similar to the initial context c_0, which increases its similarity-based recall in Equation (1).

As words are recalled, the evolution of the cue if often formalized using the AR(1) process $q_{t+1} = \rho q_t + (1-\rho)e_j$, where $(1-\rho) \in (0,1)$ is the impact of the last recollection e_j on a person's mental state (and hence on the cue). Contiguity effects arise because recall of e_j pollutes q_{t+1} with a context c_j that favors recall of the temporally contiguous words w_{j-1} or w_{j+1}, which share a similar context. Sematic proximity works analogously, via the semantic context in c_j.

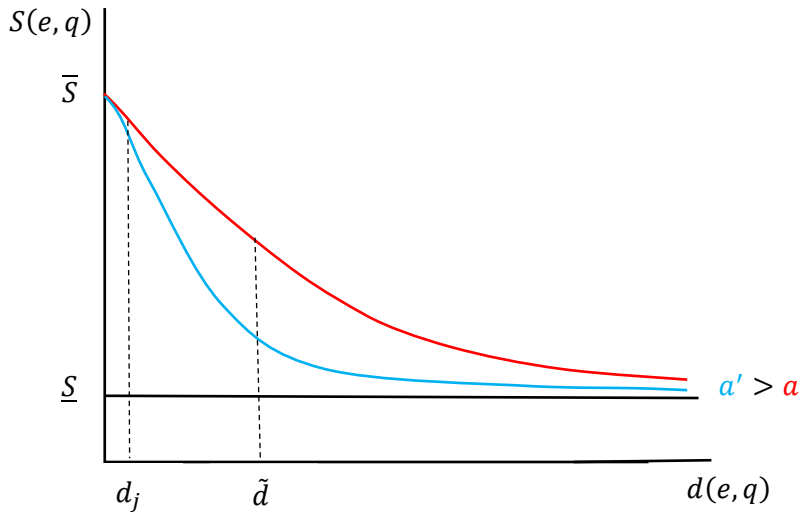
Forgetting is due to interference, the denominator in (1). In particular, recall of frequently rehearsed, or recent, or contiguous, or semantically similar words inhibits retrieval of words that do not have these properties, causing their forgetting. Critically, for the purpose of RF the most problematic kind of interference arises from irrelevant experiences: e_j can trigger recall of similar or frequent words that are *not* in the list (e.g. "apple" in the list may prompt recall of extralist "pear") or simply distract the subject, causing forgetting of list members, also reducing RF.

Recall Fluency. By Equation (1), variation in RF across people can arise from differences in the encoded frequency f_j of vectors in L and their similarity $S(e_j, q_t)$ to the cue. Such differences cannot act uniformly across experiences: interference, the driver of forgetting, is shaped by the *relative* frequency and similarity of e_j to q compared to irrelevant experiences $e \notin L$. A lower frequency f_j of trace e_j relative to the universe of extra-list frequencies $\sum_{e \notin L} f_e$ evidently reduces recall of word w_j from the list in Equation (1). Such relative frequency may vary across people, for instance because some people encode recent events more weakly (their f_j is lower) or have more numerous irrelevant experiences (their $\sum_{e \notin L} f_e$ is higher). These forces may explain why RF declines with age: older people encode fewer recent events and have more irrelevant experiences than younger ones.

Differences in perceived similarity can also play a role. They can be due to differences in peoples' attention to the features of studied words, e.g. their temporal or semantic structure, during encoding and retrieval. Attention is indeed a key driver of similarity (Nosovsky) and crucially of learning, in word tasks (Murdock, others?) and in general. Widely used memorization techniques often rely on focusing attention on contextual associations between recall targets. Such attention encodes strong associations among targets, favoring their associative retrieval and reducing interference from non-targets.

To see this formally, consider for simplicity the similarity function $S(e,q) = \mathcal{S}[a \cdot d(e,q)]$ where $\mathcal{S}$ is decreasing and $a \in (0,1)$ increases in attention to features. People with higher a exhibit a sharper similarity gradient as shown by the blue compared to the red curve in Figure X.

Figure 6. Similarity and Attention



For experiences whose features are identical to those of q, namely $d(e,q) = 0$, similarity is maximal at $\overline{S}$ regardless of attention a. Feature discrepancies instead cause attention to matter. An experience $\tilde{e}$ with discrepancy $\tilde{d}$ is perceived to be quite similar to q by a low

attention person (red curve) but dissimilar from it by a high attention person (blue curve).[7] The probability of recalling any studied word in period t from list L obeys:

$$\Pr(e_j \in L|q_t) = \frac{\sum_{e_j \in L} \mathcal{S}[a \cdot d(e_j, q_t)] \cdot f_{e_j}}{\sum_{e' \in E} \mathcal{S}[a \cdot d(e_j, q_t)] \cdot f_{e'}}. \quad (2)$$

By differentiating (2) around $a_0 = 0$ we can see that:

$$\frac{\partial \Pr(e_j \in L|q_t)}{\partial a} > 0 \quad \Leftrightarrow \quad \frac{\sum_{e_j \in L} d(e_j, q_t) \cdot f_{e_j}}{\sum_{e_j \in L} f_{e_j}} < \frac{\sum_{e \notin L} d(e, q_t) \cdot f_e}{\sum_{e \notin L} f_e}, \quad (3)$$

which links attention to RF: higher a increases the probability of recalling a target e_j if words in the list have on average more features in common with the task-cue q_t compared to other experiences. This is plausible, because items in the word list share many features: the time and space when they are studied, their phonetics, semantics, the goals of the task etc. In Figure X, then, a target word w_j at short distance d_j from the task-cue q_t exhibits less interference from $\tilde{e}$ for a person having higher attention to features, who then exhibits higher RF.

In sum, a selective recall model producing well established regularities can produce variation in RF either through variation in the encoded frequency of studied words or in attention to features. To understand the role of memory in the expectation facts of Section X, we adopt the attention channel, which we think is appropriate for two reasons. First, attention is a general-purpose factor shown to affect perception and behavior across many laboratory and field tasks, and varies systematically across people (Rosenberg Finn et al 2016, Unsworth Miller Strayer

[7] Figure 6 assumes bounded similarity, which implies that attention matters little for highly discrepant experiences, i.e. large $d(e, q)$. This property is not important for our results and can be relaxed with unbounded similarity.

2024, Hirmas Engelmann van der Weele 2024). Second, attention is parsimonious because differences in a produce effects that are isomorphic to differences in encoded frequency (e.g., low attention to current context features reduces recall of recent experiences).[8] And considering only one factor simplifies the analysis.

3.3 Selective Recall and Beliefs

The respondent forms beliefs about event $T_{\mathrm{X}} \subset \mathbb{R}$, where $X = \mathrm{SP}, \mathrm{HP}, \mathrm{LE}$ is the domain: stock market (SP), housing (HP), life expectancy (LE). For stock market and housing T_{X} is a positive price increase in the next 1 year, for life expectancy it is being alive for at least the next ten years. Following Bordalo et al. (2024), people form beliefs by: i) selectively recalling $e \in E$ based on frequency, similarity, interference, and by ii) using the recalled experiences to imagine T_{X}. The cue for person i at time t is a vector in E that specifies: the current signal $s_{\mathrm{X},it}$ (recent stock return, house price growth or personal health), its semantic context $c = \mathrm{X}$ (we abstract from temporal context here), and the respondent's affective state v_{it} (higher values stand for better mood), so $q_{\mathrm{X},it} = \left(s_{\mathrm{X},it}, v_{it}, \mathrm{X}\right)$. These features ignite selective retrieval from the database. We first focus on the role of $s_{\mathrm{X},it}$ and $c_{\mathrm{X},t}$ but valence will also play a role.

Upon retrieving $e \in E$, the DM uses it to imagine an outcome in the target space $\mathbb{R}$ based on a simulation function $\sigma_{\mathrm{X}}: E \to \mathbb{R}$ that obeys the following properties.

Assumption 1 *When forming beliefs on* $\mathrm{X} = \mathrm{SP}, \mathrm{HP}, \mathrm{LE}$*, simulation based on* $e = (s, v, c)$ *obeys:*

[8] We assume that attention to all features is a. In principle, for a total attention budget, people may vary in how they focus on one feature compared to others. This alternative formulation would not change our main results.

$$\sigma_{\mathrm{X}}(e) = \begin{cases} h(s) & if \quad c = \mathrm{X} \\ h_{\mathrm{X}}(v) & if \quad c \neq \mathrm{X} \end{cases}. \tag{4}$$

where function $h: \mathbb{R} \to \mathbb{R}$ *is increasing. Function* $h_{\mathrm{X}}: \mathbb{R} \to \mathbb{R}$ *is increasing if and only if higher values in* X *are good for the DM.* h_{X} *takes a default value* δ_{X} *for a "neutral" experience* ($v = 0$).

When the DM retrieves a Domain-Specific (DS) experience in the target domain, $c = \mathrm{X}$, she imagines the future using an increasing function h of the recalled signal s. A DM is more optimistic about stock returns if she recalls a 10% rather than a 1% stock return. People recalling higher stock, housing or health experiences make higher forecasts in the respective domain.

When the DM retrieves a Non-Domain-Specific (NDS) experience from a non-target domain, $c \neq \mathrm{X}$, she simulates based on affect, imagining rosier future if the retrieved NDS experience has higher valence v. Recalling pleasurable personal financial experiences (i.e. earning a higher salary) helps imagine high stock returns and vice versa if bad experiences are recalled. If the retrieved experience is neutral, simulation is inhibited: NDS from routine activities such as having our usual breakfast, walking to the same office, etc., act as distractors not as simulators. In this case the agent picks an agnostic default δ_{X} in the response scale such as "I do not know", the midpoint, or a random element.

In sum, simulation is more powerful when the experience that is currently top of mind is more similar to the target. It is sharper for DS experiences, it carries to NDS experience that are congruent with the current affective state, but it eventually becomes insensitive to the features of a very different experience that differ both semantically and affectively from the cue. Realistically, simulation is also less effective if the NDS comes from a domain c that is very

different from X, captured by a flatter h_{X}, as we formally and empirically show in Bordalo et al. (2026). Our predictions are strengthened if we allow for this property.

Based on recall and simulation, the future state in X imagined on average across many memory samples by an agent facing a cue $q_{\mathrm{X},it}$ and database $\mathbf{f}_i$ is equal to:

$$\mathbb{E}_{it}(s_{\mathrm{X}}) = \sum_{e:\sigma_{\mathrm{X}}(e)\in T_{\mathrm{X}}} r_e\left(q_{\mathrm{X},it}, \mathbf{f}_i\right) \cdot \sigma_{\mathrm{X}}(e). \tag{5}$$

The probability associated to the target event T_{X} then increases in such average imagination. Imagining higher future stock, prices, house prices, or a longer life leads to a higher assessed probability of positive stock or house price growth and of living for more than ten extra years.[9]

To study the implications of memory we linearize $S(e,q) = \mathcal{S}[a \cdot d(e,q)]$ at the zero attention (and hence constant similarity) benchmark $a = 0$, which gives the following result.

Proposition 1 *Let* $|\mathbf{f}_{it}|$ *be the total frequency in the database of agent* i *at time* t *and* $d\left(q_{\mathrm{X},it}\right)$ *the average discrepancy of* $e \in E$ *from the cue* $q_{\mathrm{X},it}$ *in the database. Average simulation of* X *obeys:*

$$\mathbb{E}\left(s_{\mathrm{X}}|q_{\mathrm{X},it}, \mathbf{f}_{it}\right) \approx \sum_e \frac{f_{e,it}}{|\mathbf{f}_{it}|} \cdot \sigma_{\mathrm{X}}(e) + a_i \cdot \sum_e \frac{f_{e,it}}{|\mathbf{f}_{it}|} \cdot \left[d\left(q_{\mathrm{X},it}\right) - d\left(e, q_{\mathrm{X},it}\right)\right] \cdot \sigma_{\mathrm{X}}(e). \tag{6}$$

The first term on the right-hand side says that the belief is anchored to average simulation in the database: a person is more optimistic about X (e.g. stock returns), if she had more positive DS experiences (high stock returns) or more NDS experiences that prompt stock market optimism (e.g. positive personal financial experiences) compared to an otherwise identical person.

[9] Results would change little if the assessed probability of the target event is modelled as the probability of recalling experiences that allow to simulate it, $\sum_{e:\sigma_{\mathrm{X}}(e)\in T_{\mathrm{X}}} r_e\left(q_{\mathrm{X},it}, \mathbf{f}_i\right)$, but Equation (5) is more tractable.

The second term captures our key mechanism: higher RF a_i increases the sensitivity of beliefs to experiences that share more features with the cue than average. When thinking about a domain, say the stock market, high RF reduces interference from many neutral NDS experiences whose features are incongruent with the cue. As a result, it causes stock market beliefs to be more sensitive to context-congruent stock market experiences or valence-congruent NDS ones.

We now derive specific predictions by on the role of DS and NDS experiences by assuming the following tractable discrepancy function:

$$d\left(e, q_{\mathrm{X},it}\right) = \begin{cases} \delta \cdot (s_{it} - s)^2 & if \quad c = \mathrm{X} \\ K & if \quad c \neq \mathrm{X}, v \subseteq v_{it}. \\ K + V & otherwise \end{cases} \tag{7}$$

where $\delta, K, V > 0$ and δ is sufficiently low that $\delta \cdot (s_{it} - s)^2 < K$ for realistic signal values. If experience $e = (s, v, c)$ is DS, $c = \mathrm{X}$, its discrepancy increases in the distance between its signal s and the signal in the cue s_{it}. When thinking about the stock market during a 10% return period, a 3% stock return experience is more discrepant than a 9% one. The strength of this effect increases in coefficient δ, that can capture the salience of s_{it}.

An NDS experience exhibits a discrepancy $K > 0$ due to context mismatch if it is valence congruent and an additional discrepancy $V > 0$ if it is also valence-incongruent. When thinking about stocks in a good mood ($v_{it} = \mathbb{R}_{++}$) positive valence NDS experiences exhibit lower discrepancy than negative valence ones. Both NDS experiences are however more discrepant than a DS experience due to context mismatch.

We now derive three key predictions linking RF to beliefs via selective memory. The first prediction concerns the link between RF and the prevalence of agnostic beliefs.

Prediction 1 (RF *and agnosticity*) *Let $f_{n,it}$ be the frequency of neutral experiences ($v = 0$). The probability of an agnostic default is $\frac{f_{n,it}}{|\mathbf{f}_{it}|}\{1 + a_i \cdot [d(q_{\mathrm{X},it}) - K - V]\}$, which decreases in a_i.*

Due to higher attention to features a_i, high RF people are less likely to recall irrelevant (context-incongruent and valence neutral) experiences, reducing the use of the agnostic default. This is a key mechanism: higher RF fosters active simulation through less interference.

Prediction 2 (DS, RF *and beliefs*) *Let $\overline{s}_{it}$ be the average DS signal in the database. Upon a linear approximation of $h_1(s)$ in (4), there are positive constants $\eta_0, \eta_1, \eta_2, \eta_3$ such that beliefs obey:*

$$\mathbb{E}(s_{\mathrm{X}}|q_{\mathrm{X},it}, \mathbf{f}_{it}) \approx \eta_0 + \eta_1 \cdot \overline{s}_{it} + \eta_2 \cdot a_i \cdot \overline{s}_{it} - \eta_3 \cdot a_i \cdot cov[s, (s - s_{it})^2]. \qquad (8)$$

The agent imagines a higher future state if she experienced a higher average state $\overline{s}_{it}$ in her DS experiences, $\eta_1 > 0$. This "experience effect" is entirely driven by frequency-based recall from the database. Due to similarity-based recall, though, there are two additional effects. First $\eta_2 > 0$ says that the marginal impact of $\overline{s}_{it}$ is stronger for a person with higher RF: when cued to think about, say, the stock market this person more sharply recalls DS stock market experiences, thus being more sensitive to their average value $\overline{s}_{it}$ in the database.

Second, among DS experiences, $\eta_3 > 0$ implies that higher RF fosters recall of signals s more congruent with the current signal s_{it}. This boosts beliefs if the current signal is closer to high past ones and more distant from low past ones, namely $cov[s, (s - s_{it})^2]$ is lower. Upon seeing a 10% stock return, a person is more likely to recall high experienced returns if these are close to 10%, in which case her beliefs are also more optimistic, the more so the higher is RF. Evidence for this form of cued recall has been found by Gennaioli et al. (2025) who show that the

inflation expectations of US households are more sensitive to inflation levels that are close to past experiences, as captured by a negative covariance coefficient. Our key new prediction here is that this effect should be stronger for people with higher RF, who better recall the similar past.

In sum, by creating a sharper similarity gradient, higher attention and hence higher RF increases the sensitivity of beliefs to the DS database $\overline{s}_{it}$ compared to the NDS database, and increases – within the DS database – the sensitivity to experiences similar to the current signal. We now show a third role of RF: its effect across different NDS experiences.

Prediction 3 *Suppose that NDS experiences are frequent enough relative to DS ones. Then, there is an* $\eta_4 > 0$ *such that having lived a* v*-congruent NDS experience affects beliefs as follows:*

$$\left.\frac{\partial \mathbb{E}\left(s_{\mathrm{X}} | q_{\mathrm{X},it}, \mathbf{f}_{it}\right)}{\partial f_{e,it}}\right|_{v \subseteq v_{it}} \approx \frac{1}{|\mathbf{f}_i|} \cdot h_{\mathrm{X}}(v) + \eta_4 \cdot a_i \cdot h_{\mathrm{X}}(v) \tag{9}$$

Having lived a valence congruent NDS fosters its recall and simulation, increasing the forecast if this NDS is "belief-boosting" $h_{\mathrm{X}}(v) > 0$ and decreasing them otherwise. Critically, if NDS experiences are many compared to DS ones, this effect increases with high RF. In this case, in fact, interference hinders recall of all experiences sharing any feature with the cue, including an affective feature. As a result, reducing interference via higher a_i increases the NDS effect. Thus, RF should amplify the baseline NDS effect (where the latter however can be small in a large database, when $1/|\mathbf{f}_i|$ is small).

In sum, a selective memory model of belief formation predicts a systematic relationship between RF and belief formation across domains that can be tested by studying how RF is

connected to the sensitivity of beliefs to different DS and NDS experiences. These effects are not consistent with omitted variables explanations for the role of RF in Section 2.

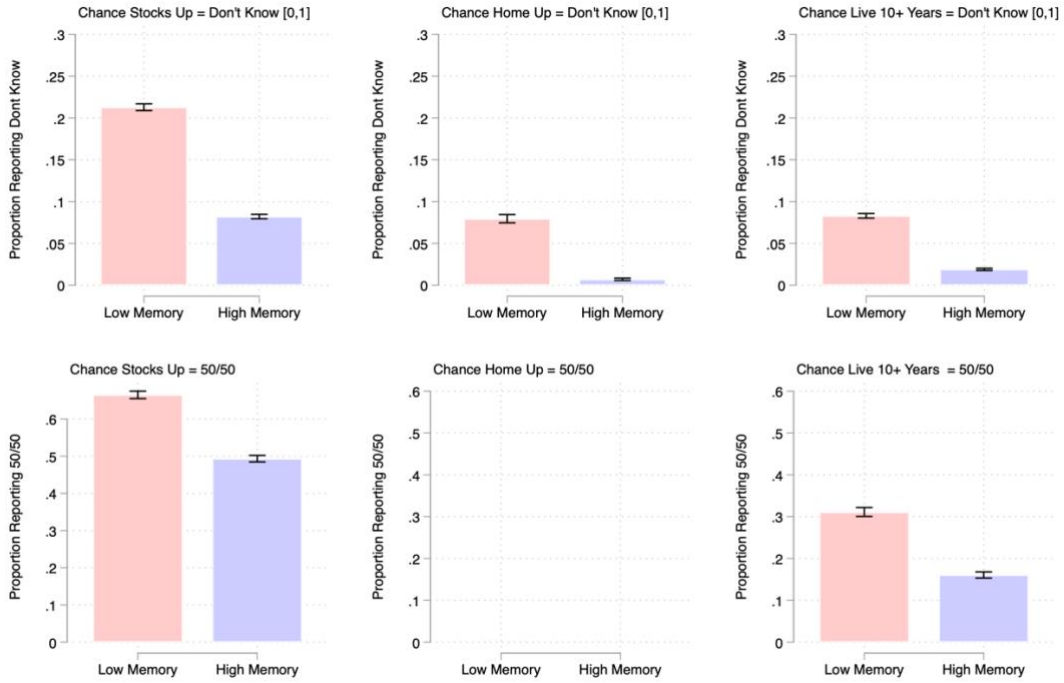
A mechanism whereby high RF proxies for better information cannot account for a heightened role of uninformative NDS experiences (but also for a stronger effect of personal DS experiences, because arguably a more informed person should also know more beyond her past experiences). Preference based explanations, according to which high RF proxies for motivated beliefs cannot explain the asymmetric role of RF in amplifying the reaction to positive and negative news for people having DS databases (the stronger impact of $cov[s, (s - s_{it})^2] \gtrless 0$). We now test the model's predictions across our three belief domains.

4.1 Agnosticity by Recall Fluency

We now test Prediction 1. Low RF people should be more likely to give agnostic answers than high RF people. This should be true in all belief domains, and arises because low RF people face more interference from neutral NDS and more generally from NDS that are not useful for simulation. Thus, they more often fail to imagine the future and hence report agnostic answers, consistent with Hassabis Shachter et al (CITE) role of memory in imagination.

We measure agnostic answers in two ways: i) share of people answering by "I do not know", and ii) share of people answering with "50-50" *and* stating that they are highly uncertain of their answer. In both cases, respondents reveal to be unsure on how to think about the future. Figure 7 below show that prediction 1 strongly holds in the data: people in the bottom RF quartile are more likely to give agnostic answers than people in the top RF quartile.

Figure 7. Probability of Giving Agnostic Answers, by Recall Fluency



Consistent with the model, we see greater prevalence of agnostic answers for low RF people. Agnostic answers are typically explained with noise: people receive a signal, are aware of its limited informativeness, and thus shrink the answer toward the prior, exhibiting low confidence. Memory explains agnostic answer by nothing coming to mind. It can potentially imply that people do not bother to shrink, but rather report "I do not know" or anchor their estimate to a salient albeit arbitrary responses. This could also explain why answers of low memory people appear more dispersed and take extreme values in Figures 3, 4 and 5.

One potential concern is that low RF people may just be less willing exert effort, both in laboratory recall tasks, where they then recall fewer words, and in forecast elicitations, where they then give agnostic low effort answers. We now test the more specific Predictions 2 and 3, which rely on the precise memory mechanism, linking past experiences and beliefs. We will also

see that memory and the report beliefs it entails shape actual economic choices, which shows that memory helps predict real world behavior, is not just a proxy for survey attention.

4.2 Domain Specific Experiences

To test Prediction 2 on the role of DS experiences we estimate Equation (8) in our three domains, stock market, housing, and life expectancy. The nature of domain specific experiences varies but the memory structure of Equation (8) and the role of recall fluency is common; The estimated probability $p_{\mathrm{X},it}$ in domain $\mathrm{X} = \mathrm{SP}, \mathrm{HP}, \mathrm{LE}$ by respondent i at time t should exhibit:

$$p_{\mathrm{X},it} \approx \gamma_0 + \gamma_1 \cdot \overline{s}_{it} + \gamma_2 \cdot a_i \cdot \overline{s}_{it} - \gamma_3 \cdot a_i \cdot cov[s, (s - s_{it})^2], \qquad (10)$$

where all coefficients are positive. What varies across domains is the computation of DS regressors. In the stock market, $\overline{s}_{it}$ is the average yearly return experienced by the respondent in the past, the observable cue s_{it} is the recent stock return that triggers selective retrieval of similar past returns s encoded in the same respondent's DS stock market memories.

In housing, $\overline{s}_{it}$ is the average house price growth experienced by the respondent in the past, the observable cue s_{it} is the recent house price growth that triggers selective retrieval of similar past house price growth s encoded in the respondent's DS housing market memories.

In life expectancy, $\overline{s}_{it}$ is the average health state experienced by the respondent in the past, the observable cue s_{it} is the recent health state that triggers selective retrieval of similar past health states s encoded in the respondent's DS health memories.

To compute these proxies, for each respondent we construct a memory database of: 12 month stock returns and house price growth at the quarterly frequency since the date of birth,

and the measured change in the personal health PCA in each HRS wave.[10] For the latter we have relatively few observations, equal to the number of waves a respondent participates to. We however measure health DS also by childhood health, which we include as a regressor. In stock market and house price regressions we can directly control for long term health outcomes using the stock or housing returns experiences throughout one's life.

Table 1 below reports, for each domain, the estimation of the model Equation (10) in columns 1,3, 5. In the columns 2, 4 and 6 we add RF as a control to check that RF does not merely capture the level effect on beliefs but rather a sensitivity to past DS experiences. In the Appendix, we show additional specifications that include i) year fixed effects to show that results are not driven by common shocks but rather specific memory reactions, and ii) the interaction between RF and IQ, to show that effects are not driven by respondents having different information sets.

The Table reveals the same general structure for beliefs in different domains, and strongly supports the role memory. In each domain: beliefs increase in "positive past DS", but especially so for high RF people. Cues also matter, as shown by the negative coefficient on the covariance term. The same news exhibits a stronger effect for people who have high RF and live through similar DS experiences. Similarity matters and works through RF.

[10] The results do not change if, following Gennaioli et al (2025) we use experiences starting at age 16. See Appendix.

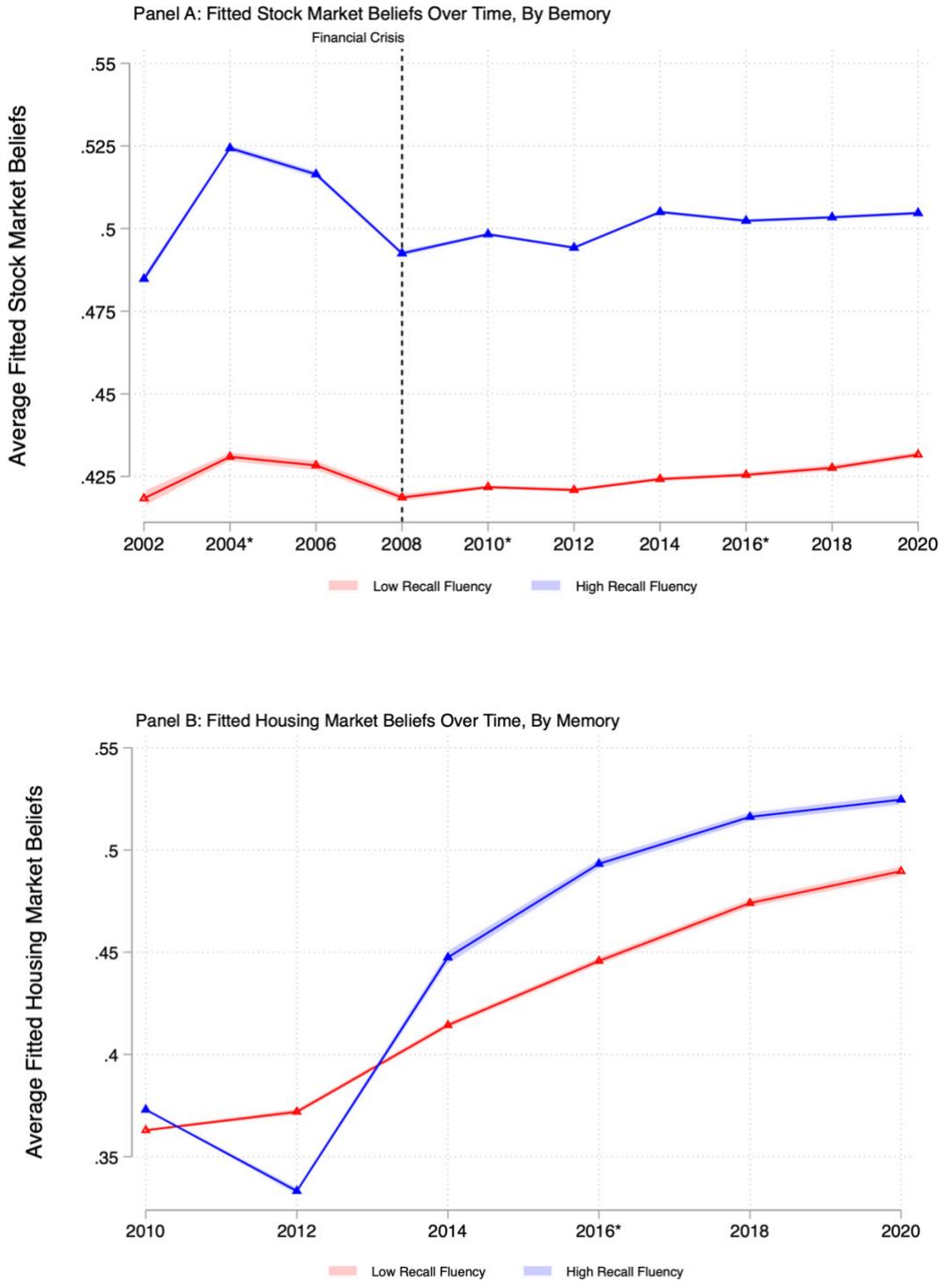
Table 1. Beliefs, RF and DS

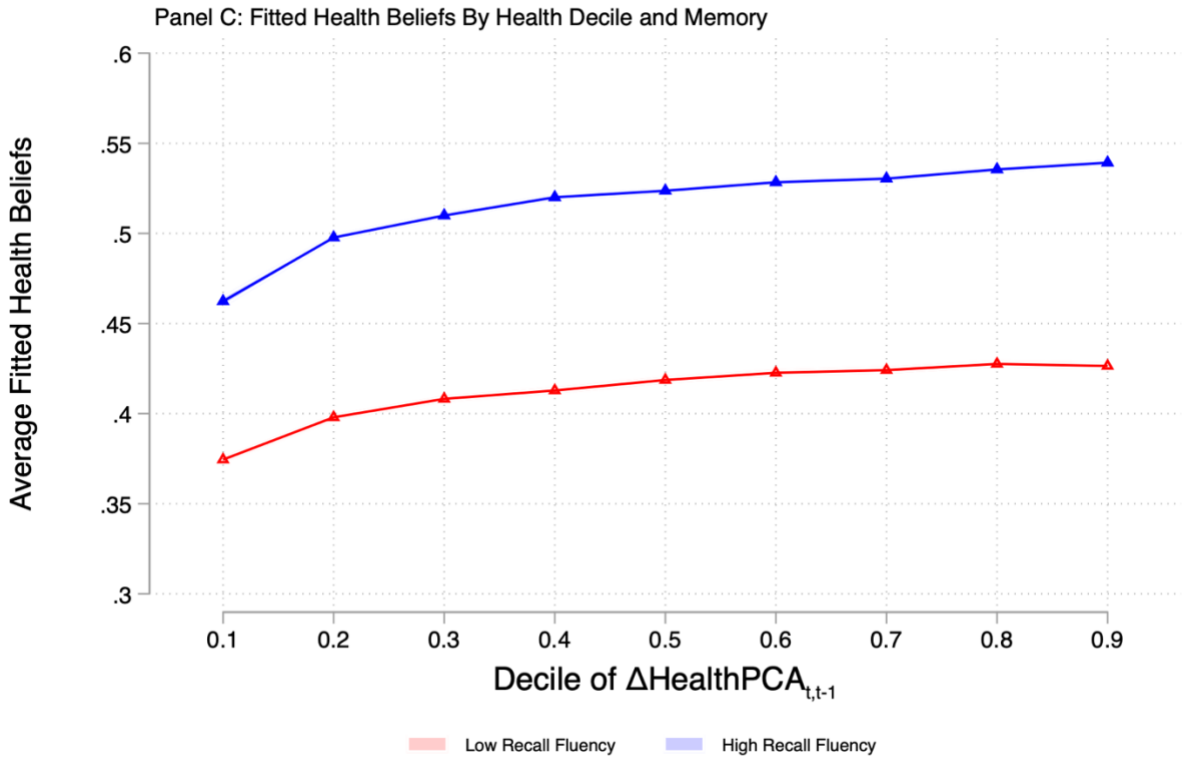
	(1)	(2)	(3)	(4)	(5)	(6)
	Chance Stocks ↑		**Chance House ↑**		**Chance Live10+**	
Cognition						
RF	.	0.001	.	-0.010	.	0.014***
	.	[0.004]	.	[0.008]	.	[0.001]
IQ	0.014***	0.014***	0.009***	0.009***	-0.007***	-0.010***
	[0.001]	[0.001]	[0.002]	[0.002]	[0.001]	[0.001]
DS Experiences						
$\overline{s_{it}}$ (γ_1)	0.000	0.001	0.022***	0.017***	0.033***	0.031***
	[0.001]	[0.002]	[0.002]	[0.005]	[0.002]	[0.002]
RF Interactions						
$RF * \overline{s_{it}}$ (γ_2)	0.017***	0.015***	0.016***	0.028***	0.010***	0.011***
	[0.001]	[0.005]	[0.002]	[0.010]	[0.002]	[0.002]
$RF * cov(s, (s - r_t)^2)$ (γ_3)	-0.007***	-0.007***	-0.065***	-0.065***	-0.011***	-0.011***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
N	139345	139345	49972	49972	129518	129518
AR2	0.04	0.04	0.05	0.05	0.10	0.10
Demographic Controls	Y	Y	Y	Y	Y	Y
Cohort FE	N	N	N	N	Y	Y
Age FE	N	N	N	N	Y	Y

Table 1 shows that memory selectivity of experiences is a key technology of beliefs. Recall fluency is quantitatively central to understand heterogeneity in beliefs and in the reaction to information, the key drivers of beliefs in the field, as shown in the next Figure.

Finally, Figure 9 shows how the model of Table 1 reproduces the main empirical patterns of beliefs about stocks, housing, and life expectancy described in Section 2, including the higher volatility and lower heterogeneity of beliefs held by high RF respondents.

Figure 8. Fitted Beliefs Over Time, By Memory





4.3 Non Domain Specific Experiences

A distinct implication of memory is that beliefs may be shaped by experiences that are not domain specific, see Prediction 3. To test this prediction we add, for each domain, proxies for NDS and their interaction with RF. For stocks and housing we use childhood health and childhood socioeconomic status. These are remote and entirely personal experiences which have no relevance and contain no information for stock and housing markets. For life expectancy we use experienced stock and housing returns, which are also from a different domain and broadly not informative about health.[11] These proxies are included alone and interacted with RF.

[11] Discuss wealth effects.

Table 2. NDS, RF and Beliefs

	(1)	(2)	(3)	(4)	(5)	(6)
	Chance Stocks ↑		**Chance House ↑**		**Chance Live 10+**	
Cognition						
IQ	0.014***	0.013***	0.009**	0.009***	-0.010***	-0.012***
	[0.001]	[0.001]	[0.003]	[0.002]	[0.001]	[0.001]
DS Experiences						
$\overline{s_{it}}$ (γ_1)	0.000	0.001	0.022***	0.023***	0.031***	0.031***
	[0.001]	[0.001]	[0.006]	[0.002]	[0.002]	[0.002]
RF Interactions						
$RF * \overline{s_{it}}$ (γ_2)	0.017***	0.016***	0.016***	0.015***	0.011***	0.012***
	[0.001]	[0.001]	[0.003]	[0.002]	[0.002]	[0.002]
$RF * cov(s_k, (s_k - r_t)^2)$ (γ_3)	-0.007***	-0.007***	-0.065***	-0.065***	-0.011***	-0.012***
	[0.001]	[0.001]	[0.018]	[0.001]	[0.001]	[0.001]
NDS Experiences						
nds (δ_1)	-0.001	-0.003	-0.005	-0.006	-0.001	0.001
	[0.002]	[0.002]	[0.004]	[0.004]	[0.002]	[0.001]
$RF * nds$ (δ_2)	0.004**	0.007***	0.006*	0.009**	0.015***	0.017***
	[0.002]	[0.002]	[0.003]	[0.004]	[0.001]	[0.001]
N	134136	139342	49593	49971	129518	111536
AR2	0.04	0.04	0.05	0.05	0.10	0.11
NDS Var	Chld.Health	Chld.SES	Chld.Health	Chld.SES	$\Delta_L sp500$	$\Delta_L hpi$
Demographic Controls	Y	Y	Y	Y	Y	Y
Cohort FE	N	N	N	N	Y	Y
Age FE	N	N	N	N	Y	Y

Table 2 reports the results. NDS experiences matter for beliefs in all domains, but especially – or almost exclusively – for high RF people. When these people form beliefs, cues in the environment such as their current mood prompt them to recall NDS experiences which affects simulation. Good childhood health and socioeconomic status predict more optimistic about positive stock returns and house price growth (these are homeowners, so they like the house price growth). Good economic return predicts more optimism about life expectancy.

The fact that NDS only matter for high RF people is consistent with the mechanism whereby these events matter through memory, not because they capture omitted variables. This

finding also suggests that high RF people are not just more intelligent and better informed. They use memory more, including the normatively irrelevant NDS (especially in the case of macro outcomes).

5 Recall Fluency, Expectations and Choices

Does memory affect economic choice by shaping, through recall fluency, the role of DS and NDS experiences? To address this question, we first correlate beliefs with choice. Next, we perform a two-stage exercise assessing the role of RF and experiences. In the first stage, we regress a respondent's choice on a proxy for beliefs about a variable relevant to that choice:

$$C_{it} = \propto + \eta^M RF_{it} + \eta^I IQ_{it} + \theta B_{it} + \beta X_{it} + \varepsilon_{it}, \quad (3)$$

where C_{it} is a choice recorded for respondent i at time t (such as new stock market investments), and B_{it} is the measured choice-relevant belief (such as belief about stock returns) of the same subject and the same time.

In the next step, we study whether, by shaping beliefs, baseline experiences and recall fluency matter for choice. To do so, we take the belief specifications of Table 2 and we separate fitted beliefs into three components: $\widehat{B_{it}} = B(E_{it}) + B(RF_{it} \cdot E_{it}) + B(Dems_{it})$. The first term $B(E_{it})$ is a pure "Experience" component, constructed by multiplying the respondent's experience of both DS and NDS by their respective estimated coefficients from Table 2.[12] The second term, $B(RF_{it} \cdot E_{it})$, is an "Amplification" effect, constructed by multiplying the

[12] Formally, $B(E_{it}) = \widehat{\gamma_1} \cdot \overline{s}_{it} + \widehat{\delta_{1,health}} \cdot ChildHealth + \widehat{\delta_{1,ses}} \cdot ChildSES$. Instead of estimating the impact of childhood health and childhood SES separately as we do in Table 2, we estimate $\widehat{\delta_{1,health}}$ and $\widehat{\delta_{1,ses}}$ jointly in regression that includes both regressors.

respondent's level $RF_{it} \cdot E_{it}$ by its estimated coefficient from Table 2.[13] This generates the part of beliefs that are predicted by memory's amplification of experiences. Finally, we take the remaining *fitted* belief, $B(Dems_{it})$, as beliefs predicted from demographic variables and IQ. Using all three components, we then estimate the following specification:

$$C_{it} = \propto + \eta^M RF_{it} + \eta^I IQ_{it} + \gamma B(E_{it}) + \mu B(RF_{it} \cdot E_{it}) + \rho B(Dems_{it}) + \beta X_{it} + \varepsilon_{it}, \qquad (4)$$

where $\gamma > 0$ means that greater exposure to experiences that drive belief B_{it} also promotes the choice itself, while $\mu > 0$ means that this is especially true for people having high recall fluency. We measure choice and choice relevant beliefs as follows.

Stock Market. Choice variables are: *New Stock Market Investments 2y*, a dummy capturing whether the respondent has made any new stock market investments in the last 2 years; *% IRA in Stocks,* measuring the proportion of the respondents IRA that is invested in stocks/mutual funds (this variable is restricted to people that own an IRA). The relevant belief B_{it} is the assessed probability of stocks going up, which should promote both measures of stock investment.

Housing. The choice variable is *Bought New Home 2y*, a dummy taking value one if the respondent bought a new home in the past two years. The relevant belief measure is the probability of house prices going up, which should promote this choice measure.

Life Expectancy. Choice variables include: *Chance Work Past 70*, a variable from 0 to 100 indicating the believed probability for the respondent to work after age 70 (the variable is restricted to people who are currently working and younger than 70); *Retired*, a dummy taking

[13] Formally, $B(RF_{it} \cdot E_{it}) = \widehat{\gamma_2} \cdot a_i \cdot \overline{s}_{it} + \widehat{\delta_{2,health}} \cdot a_i \cdot ChildHealth + \widehat{\boldsymbol{\delta}_{2,ses}} \cdot a_i \cdot ChildSES.$

value one if the respondent is currently retired. Importantly, in regressions where the choice variable and belief are related to health, such as the chance to work past 70 and whether the person is retired, we include age fixed effects; *% Income to Pension*, namely the percentage of monthly pay the respondent contributes to their pension (restricted to people who are currently working and have a defined contribution pension scheme). The belief measure relevant to these choices is the probability attached to being alive in the next ten years, which should be associated with higher believed probability of working after 70, a lower probability of being currently retired, and a higher contribution to pension. In these regressions we control for a range of individual characteristics which may affect choices independently of beliefs, such as age, education, gender, income, and marital status.

Table 3 reports, for each choice, the estimates specification (3) in the first column and the specification (4) assessing the role of experiences and recall fluency in the second column. In this latter specification we also include a preference proxy, namely the self-reported degree of risk aversion. In economics risk aversion is considered a key proxy for interpersonal differences in behavior so including this proxy allows us to benchmark the explanatory role of memory against standard preference measures. To measure risk aversion, we take the standardized average of 3 proxies: a) risk taking for employment, which asks the respondent if they would prefer a job with higher expected pay but more variance vs one with lower pay and more certainty, b) risk taking for business inheritance, which asks the respondent if they would put an inheritance into a safe asset or invest in a risky business, and c) self-reported risk taking, measured on a scale of 1 to 5.

Table 3. Beliefs, RF, and Choices

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Stock/Housing Market Choices					
	New Stock Inv. 2y (0,1)		*Pct. IRA in Stocks (0-1)*		*Bought New Home (0,1)*	
Cognition						
IQ (η^I)	0.005***	-0.004***	0.027***	0.007*	0.005***	-0.010***
	[0.001]	[0.001]	[0.004]	[0.004]	[0.001]	[0.004]
Beliefs						
$\widehat{B_{it}}$ (θ)	0.062***	.	0.076***	.	0.007***	.
	[0.003]	.	[0.007]	.	[0.001]	.
$\widehat{B(E_{it})}$ (γ)	.	0.007***	.	0.008***	.	0.003***
	.	[0.001]	.	[0.002]	.	[0.001]
$\widehat{B(E_{it} \cdot RF_{it})}$ (μ)	.	0.019***	.	0.020***	.	0.004***
	.	[0.001]	.	[0.002]	.	[0.001]
Risk Taking						
RiskTaking	.	0.003***	.	0.027***	.	0.004***
	.	[0.001]	.	[0.003]	.	[0.001]
Avg DV	0.10	0.10	0.56	0.56	0.04	0.04
N	129529	120357	36078	34263	45249	43865
AR2	0.06	0.06	0.06	0.06	0.01	0.01
Beliefs About	stocks	stocks	stocks	stocks	house	house
	Panel B: Health Choices					
	Chance Work Past 70 (0-1)		*Retired (0,1)*		*% Inc. to Pension (0-1)*	
Cognition						
IQ (η^I)	0.020***	0.016***	-0.007***	0.000	0.002	0.002*
	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]
Beliefs						
$\widehat{B_{it}}$ (θ)	0.030***	.	-0.013***	.	0.003*	.
	[0.003]	.	[0.002]	.	[0.002]	.
$\widehat{B(E_{it})}$ (γ)	.	0.014***	.	-0.011***	.	0.002
	.	[0.002]	.	[0.001]	.	[0.001]
$\widehat{B(E_{it} \cdot RF_{it})}$ (μ)	.	0.014***	.	-0.008***	.	0.001
	.	[0.002]	.	[0.001]	.	[0.001]
Risk Taking						
RiskTaking	.	0.028***	.	0.002	.	0.002
	.	[0.002]	.	[0.001]	.	[0.001]
Avg DV	0.26	0.26	0.60	0.60	0.08	0.08
N	26078	25723	111536	101636	5707	5645
AR2	0.05	0.06	0.42	0.42	0.02	0.02
Cohort Fixed Effects	Y	Y	Y	Y	Y	Y
Age Fixed Effects	Y	Y	Y	Y	Y	Y
Beliefs About	health	health	health	health	health	health

Table 3 reports the results. Across all domains, choices are driven by beliefs (columns 1,3, and 5). While beliefs driven by experiences $\widehat{B(E_{it})}$ matter in their own right, beliefs generated

from memory's amplification of experiences $B(\widehat{RF_{it}} \cdot E_{it})$ tend to have a larger impact. This suggests that the way in which memory retrieves and weights past experiences matters more than the experiences themselves.